Monica J. McCormick. Filling Institutional Repositories by Serving the University's Needs. A Master's paper for the M.S. in L.S. degree. April, 2006. 76 pages. Advisor: Jeffrey Pomerantz

Institutional repositories (IRs) are digital archives of university-owned and -created content. They are a solution that librarians have proposed to address the high costs and limited library control of much published digital scholarship. IRs are one facet of the open access movement, designed to make it easier for libraries to obtain, organize, and preserve their universities' digital assets, and offer them freely on the internet. The long-term goal is to create an alternative to traditional publishing. Despite significant developments in creating software and crafting metadata and preservation standards, the institutional repositories opened so far have not attracted significant amounts of content.

This paper suggests that institutional repositories will more successfully challenge the current system of scholarly communication if they first address the needs of local stakeholders: not only the library, but also the university archives, the faculty, students, information technology department, university press, and the campus administration. After describing the contexts of IR emergence, the paper examines the vision for and current deployment of institutional repositories. Finally, it explores the needs of each stakeholder group in relation to digital material, and outlines how an IR might benefit each of them. The thesis is that institutional repositories will become a strong part of the campus infrastructure only if they solve problems for stakeholders beyond the library. Once that is accomplished, we may begin to see how IRs can influence the wider system of scholarly communications.

Headings:

Institutional repositories

College and university libraries

Digital Preservation

Open Access

FILLING INSTITUTIONAL REPOSITORIES BY SERVING THE UNIVERSITY'S NEEDS

by
Monica McCormick

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

April 2006

Approved by:

_____

Jeffrey Pomerantz

INTRODUCTION

Institutional repositories have become a hot topic in academic libraries in the past few years, discussed mainly as a possible solution to some large and thorny problems of scholarly communications. In the face of rising costs, flat budgets, and restricted access to information, as well as rapid changes in technology, scholarly practice, and patron expectations, research libraries have been challenged to maintain their depth of collections and high level of service. An institutional repository (IR) is a digital archive where a university community's intellectual work is freely available and preserved for posterity. (This term will be defined in more detail below.) The concept of institutional repositories suggests the tantalizing possibility of greater library influence over the full cycle of scholarly communication on campus, from research through publication, collection, and preservation. IRs offer the added boon of opening the results to the world through the internet. Research libraries have embraced this possibility, and institutional repositories are being created. This early stage of development is characterized by passionate advocacy for open access in the library community, grant-funded efforts to design open-source software and create digital preservation standards, reactions from publishers that vary from defensive to encouraging, and, through it all, the regular opening of new digital archives large and small throughout the world. Many scholars, however, appear not to be engaged in this movement, if we are to judge by the amount of content submitted to

repositories so far. And the long-term goal of altering scholarly communications remains elusive.

This paper will examine institutional repositories—the context of their emergence, the hopes and expectations of their developers, and the extent of their implementation. Further, I will suggest that successful development of institutional repositories will require more than rhetoric, grant funding, good software, and thoughtful standards. If institutional repositories are to create new possibilities for scholarly communications from within the university, then we need to understand the complexities of that system within the university and beyond it. Librarians must consider what their campus colleagues stand to gain, and why they would invest their time and energy in this particular solution. I propose that institutional repositories will work best not as the pre-determined solution to a problem defined by librarians, but as a set of services that solve problems for groups outside the library. By examining the needs of several university stakeholders, I hope to suggest some useful avenues toward the goal of more scholarship, freely available and thoughtfully preserved.

Perhaps the first IR in the US was opened at MIT in 2002. With only a few years of existence, institutional repositories are just beginning to take shape. Institutional repositories are most simply defined as repositories of digital content, accessed via the web and supported by an institution (for the purposes of this paper, a university), to house the work of its members. The possible range of content extends to many formats–text, images, video, audio, databases, and more. The general idea is to store, manage, and

preserve a university's born-digital and digitized assets, making them freely available via the internet.

In the past half-decade, we have seen a rapid deployment of IRs on university campuses in the United States and around the world. The Coalition for Networked Resources suggests that they are "becoming well-established as campus infrastructure components." (Lynch & Lippincott, 2005) Several different types of open-source and commercial software have been developed to support them – Eprints, DSpace, Greenstone Fedora, and BePrints. These have been built with university, foundation, and corporate funding (University of Southampton for Eprints; MIT and Hewlett-Packard for DSpace; the Mellon Foundation and University of Virginia and Cornell for Fedora; University of California for BePrints.) Although it is impossible to determine precisely how many institutional repositories have been created, a Coalition for Networked Information survey of 124 research universities in the US found that 40% of respondents have some type of institutional repository operating, and 88% of those that do not are planning to implement one.(Lynch & Lippincott, 2005)

Despite this, there is evidence that it is difficult to populate IRs with content. A 2004 study found the average number of records to be only 290 per repository. (N. F. Foster & Gibbons, 2005) As with the number of repositories, it is practically impossible to count the objects in them, since those objects are defined in many ways. But the impression that attracting IR content is supported in the literature, with articles explicitly acknowledging the challenge of attracting content and responses to the problem ranging from hiring

marketing consultants to performing copyright clearance and other services. (Blummer, 2005) (M. Mackie, 2004a) This topic will be covered in more detail in the next section, but for now suffice it to say that there are evidently barriers—practical, cultural, economic—to integrating IRs into the cycle of scholarly production and consumption.

Part one of this paper explores the context in which institutional repositories have become both possible and potentially powerful. Part two examines the specific ideas about how IRs can work, and outlines their development so far. Part three suggests an approach to building a successful institutional repository by assessing the needs of important stakeholders. The discussion is in the broad context of the dysfunctions in the international system of scholarly communication. As one observer has argued, "the scholarly communication system must be re-engineered from a holistic point of view, taking into account the needs of scholars, libraries, publishers, and others with a stake in the production and circulation of knowledge." (Houghton, with Steele, & Henty, 2003) I advocate that holistic approach to building IRs. While librarians may be persuaded of the value of a locally-maintained archive of digital content, other parties on many campuses evidently do not yet see that as a priority. How to persuade them that this new service is worth investing in? Starting from the perspective of the university library, I will explore the needs of their various potential partners in this endeavor. My audience is primarily librarians and their current collaborators on IR projects, and only secondarily those stakeholders with whom they hope eventually to work. I believe that if institutional repositories can effectively addresses the needs of all relevant campus stakeholders they are far more likely to become key nodes in the web of scholarly communication.

PART ONE: CONTEXTS, RESPONSES, AND A PROPOSED SOLUTION

Context: Restricted access to information

The recent economic history of research libraries has been one of steadily rising prices and flat budgets. The Association for Research Libraries has reported that among their member libraries in the period 1986-2003, the price per subscription of serials rose by 215%, while the Consumer Price Index rose only 68%. Member libraries paid 260% more for their serial subscriptions in 2003 than in 1986 despite having increased the number of subscriptions by only 14%. (Panitch & Michalak, 2005) Concurrently, library investments in technology have climbed, as catalogs, reference tools, and content moved online, requiring new equipment and technical services. In these circumstances, most libraries have cut back on acquisitions, regularly canceling subscriptions and buying fewer monographs. This is not to suggest that libraries have been passive. But despite efforts to build collaborative collections and create consortia for more cost-effective services, in addition to building systems to enable increased access to library materials, the hard fact remains that unless budgets change dramatically and unexpectedly, research libraries may have less content on offer.

Along with the budget problems has come a loss of control over the materials for which libraries pay so dearly. With serial publications increasingly available in digital format and patrons preferring online access, many libraries have discontinued print subscriptions. This has many consequences beyond the obvious one of higher prices, not least being the organizational changes in the library's operational costs associated with shifts in staffing, resources, materials, space and equipment. (Montgomery, 2000) Of

most concern for the purposes of this paper, however, are shifts in the library's ability to store and deliver digital material over time. Digital content is licensed, not bought. Libraries cannot guarantee long-term access to material that they do not own. Furthermore, licenses may restrict how digital content is used. Licenses carry more legal weight than copyright law, so that fair uses of licensed content for educational purposes may be diminished.

A related loss of control over scholarly content takes place earlier in the cycle of scholarly communication. Academic authors seeking tenure and promotion must publish to demonstrate their academic contributions—that is, they publish for influence rather than income, generally giving publishers the copyrights for scholarly works for little or no cost. Universities, then, find themselves in the position of paying twice or even three times for the same material – in research support to the authors, license fees for the published material, and sometimes again in the price of digital backfiles of serials (bought to ensure the possibility of long-term access and preservation.)

In these circumstances, research libraries find it increasingly difficult to fulfill their traditional mission of providing broad, long-term access to scholarly content for their faculty and students.

Context: changing patterns in the production of knowledge

The development of institutional repositories also responds to the recognition that, despite continued reliance on traditional publishing venues, scholarly communication is

being conducted in more socially distributed ways. This change has been described as a shift in modes:

> In Mode 1 problems are set and solved in a context governed by the, largely academic, interests of a specific community. By contrast, Mode 2 knowledge is carried out in a context of application. Mode 1 is disciplinary while Mode 2 is transdisciplinary. Mode 1 is characterised by homogeneity, Mode 2 by heterogeneity. Organisationally, Mode 1 is hierarchical and tends to preserve its form, while Mode 2 is more heterarchical and transient. Each employs a different type of quality control. In comparison with mode 1, Mode 2 is more socially accountable and reflexive. It includes a wider, more temporary and heterogeneous set of practitioners, collaborating on a problem defined in a specific and localized context. (Gibbons et al., 1994)

Communication technology is crucial to Mode 2:

> Socially distributed knowledge production is tending towards the form of a global web whose numbers of inter-connections are being continuously expanded by the creation of new sites of production. As a consequence, in Mode 2 communications are crucial. . . . To function, the new mode needs to be supported by the latest that telecommunications and computer technologies have to offer. Mode 2, then, is both a cause and a consumer of innovations which enhance the flow and transformation of information. (Gibbons et al., 1994)

This analysis, already a dozen years old, remains relevant. New forms of collaboration take advantage of new technologies and enable further possibilities for exchanging information and creating new scholarship. The general notion described as Mode 2 is a more fluid, open, and flexible system of scholarly production, aimed at solving problems in a collaborative fashion. While the transition to this mode is by no means universal, there is evidence for it in the growing numbers of interdisciplinary research units and publications in many university settings, and in the online sharing of research results and data.

Context: Open Access

The more open circulation of scholarship has been an important underpinning of the development of institutional repositories. Perhaps it was inevitable that the question arose: why are publishers necessary? As published content grew more expensive and restricted, and the internet made the distribution of ideas relatively cheap and easy, avoiding the publisher as "middle man" became an obvious option.

An early instance of these ideas was distributed by the Association for Research Libraries in 1995. The publication reproduces an internet discussion begun by scientist Steven Harnad in a post titled, "The Subversive Proposal." Harnad called for authors of "esoteric" (i.e. specialized) work to deposit it on internet ftp servers. He said, "As soon as all research authors publicly self-archive their refereed and unrefereed papers publicly online, the research literature will be free for all."(S. Harnad, 1995) Harnad's shot over the bows engaged many other scholars, publishers, and librarians. The conversation has evolved into what is now known as the open-access movement. Open access literature is "digital, online, free of charge, and free of most copyright and licensing restrictions. What makes it possible is the internet and the consent of the author or copyright-holder. . . . There are two primary vehicles for delivering OA to research articles: OA journals and OA archives or repositories."(Suber, 2004)

Open access repositories were first created to circulate scholarship in particular disciplines. The most prominent examples are arXiv.org (for physics, math, and computer science, at http://arXiv.org) and cogprints.org (for psychology, linguistics, neuroscience,

philosophy, at http://cogprints.org). These services, noted in Harnad's subversive proposal, were founded in the early 1990s with university technical support—arXiv in the US, cogprints in the UK. Scholars posted their work pre-publication, making it available more quickly than publishers could, and allowing early commentary from colleagues. These repositories now contain pre- and post-print works, demonstrating that scholarship can be self-published and informally peer-reviewed without precluding later traditional publication. Many advocates of institutional repositories consider the disciplinary repositories a valuable model for more open access to research.

Open access journals vary widely—some are peer-reviewed, others not—but by definition they are made available online without subscription fees. The Lund University library in Sweden maintains a directory of open-access journals at http://www.doaj.org/, which as of this writing included over 2,100 journals. There are well-supported efforts to enable scholarly OA journals. For example, SPARC, the Scholarly Publishing and Academic Resources Coalition, was established by the Association for Research Libraries to "unleash the power of the digital networked environment in order to enhance the process of scholarly communication and address the serious economic problems that plague it." (Johnson, 2002) SPARC first created partnerships to support lower-cost journals "as an alternative for academic disciplines formerly dependent on high-priced journals" and now also supports open-access publishing. A more recent example is the Public Library of Science (www.plos.org), founded in 2002 to make scientific knowledge more widely available than conventional publishers were willing to do. PLoS journals are a "nonprofit scientific and medical publishing venture that provides scientists and

physicians with high-quality, high-profile journals in which to publish their most important work." Costs are supported by grants and author contributions (which generally come from research funds). (Public Library of Science)

Context: Technological developments

Beyond the academy there are other pertinent developments. One is the way that internet users, particularly young ones, behave online. On widely popular sites such as Facebook and Myspace, students exchange a wide array of personal information and connect to one another in ever-expanding social networks. Through these and other venues like Flickr and del.icio.us, net denizens create and share photos, videos, and other content, tagging items with "folksonomy" terms that enable other users to find and sort information to suit their needs. Yet another example is Wikipedia.org, a reference tool written and revised by over one million users. These are all user-made collections and services that operate according to rules much more casual than those of traditional publishers and libraries. Although the quality of such content is debated, such sites demonstrate how millions of internet users now engage, shape, and reuse online content for their own purposes.

Library Responses

What are the implications for libraries? As more scholars take part in Mode 2, more scholarship is produced in new forms and circulated in new venues, and more patrons become accustomed to organizing online information according to their personal desires,

libraries consider how to respond to new communication networks and changing patron expectations.

The concept of Library 2.0 is one way of describing how libraries conceptualize their response. This idea derives from the related notion of Web 2.0 – an approach to web services that Tim O'Reilly has defined as follows:

> Web 2.0 is the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an "architecture of participation," and going beyond the page metaphor of Web 1.0 to deliver rich user experiences. (Cited in Miller, 2005)

Library 2.0 similarly suggests a reliance on network technologies to provide dynamic content in many formats that can be used for varying purposes by a single patron or a group, regardless of physical location. Library services may be shared or outsourced to some external partners, such as a consortium (e.g., digital reference) or a community of users (e.g., collaborative tagging).

Library 2.0 is being developed in a myriad of ways. Information Commons are among the earliest developments. They are at once physical and virtual collaborative spaces – library areas provided with many computers, on which patrons can find a vast array of shared digital content and use it with a range of technology tools. (Kranich, 2004) Another example is the NC State University Libraries' OPAC, with enhanced search functions similar to an online shopping site (see http://www.lib.ncsu.edu/catalog/).

Other libraries are developing metasearch tools that link from search results directly to content, or providing MyLibrary portals so that patrons can customize their online experience. The Ann Arbor Public Library site is, at first look, a blog, from which users can find the catalog and other library services (see http://www.aadl.org/). Other technical innovations have been proposed, including reference blogs, annotatable OPACs, and more.

Libraries are still a long way from seamless access to information, since much of their content is divided into silos separated by format, rights restrictions and technological barriers. Library 2.0 is an effort to break down those walls. While this may simply be a trendy term for the familiar concept of user-focused service, its origins in the vocabulary of the Web reveal an important underpinning. Libraries exist in a rapidly changing information culture. Today's patrons employ search engines and online retail spaces that make search and retrieval quick and intuitive. Library resources are contained in more complex and rigid structures, and generally require training for optimal use. Library 2.0 encourages the thoughtful appropriation of information technologies for new library uses, hence integrating libraries more thoroughly into the larger information world.

A Proposed Solution: Institutional Repositories

In these contexts—restricted access to knowledge, changing scholarly practices, support for open access, rapidly evolving information technology—institutional repositories have emerged as a potential solution.

> The institutional repository (IR) concept has gained momentum as universities begin to question the logic of buying back [their] research, as

> libraries drop journal subscriptions due to publisher fees outstripping resources, and as taxpayers question paying for research twice by funding the research itself followed by purchasing journal subscriptions to discover the research findings. IRs can preserve and provide access to a university's unpublished material, establish alternatives to the high costs of traditional publications, and contribute to a university's prestige. As information and knowledge resources are increasingly digitized and distributed by local and global networks, those facing the above issues are exploring alternatives to the preservation and distribution of information. (McClendon, 2005)

Despite McClendon's invocation of universities, libraries, and taxpayers as builders of momentum for IRs, in fact it has been mainly librarians who have embraced the concept. And that may be one reason for their rather widespread existence but relatively sparse content: IRs are a solution to problems that the librarians see more clearly than their colleagues.

This paper, then, discusses the IR as a product or service that may need to be promoted to the campus community. The marketing literature provides one approach. It is a commonplace for cultural organizations to recognize that they must promote their services and establish brand identities. (Kotler & Levy, 1969) For institutional repositories, it is perhaps more important to consider the IR in the realm of product development. A recent Harvard Business Review article argues that in creating a new product it is important to understand the tasks your customers want to accomplish. (Christensen, Cook, & Hall, 2005)

The authors believe that successful products must be developed to solve customers' problems rather than to address the expected needs of a particular market segment. One

example illustrates the point. Seeking to increase milkshake sales, a fast-food company interviewed customers they had identified as key to the “milkshake demographic.” Based on that research, they changed their recipes, but sales remained about the same. Further research showed that 40% milkshakes were purchased early in the morning. They learned that customers facing a long commute wanted a breakfast that would be both filling and easy to eat in the car. Bananas would not hold them until lunch; bagels and breakfast sandwiches were too messy. Milkshakes, despite their relative lack of nutritional value, served these consumers’ needs. In Christensen’s language, the customers “hired” milkshakes to do a particular job. The company responded by providing milkshake dispensers in front of their counters, where customers could buy them with a simple swipe of a credit card. , and created new flavors with chunks of fruit, making the product more fun to eat. Milkshake sales improved. (Christensen et al., 2005)

The point of this story is that creators of a new product must ensure that is does what their customers need—and the needs assessment may reveal some surprises. I will apply this insight to investigate how institutional repositories might solve problems: what do various campus groups want and need to do with digital assets? Can an institutional repository accomplish those tasks? How?

The groups I consider are the library, university archive, faculty, students, information technology managers, university press, and campus administration. All of them have a stake in what is done with digital assets, but may approach their care and handling from different perspectives. I assume that a successful institutional repository

may require the support of all of these groups, and therefore must serve them. I will examine each group's needs in relation to digital content. I will also explore the benefits that each might receive from an IR and how they might contribute to its success. My aim is to explore how university libraries, through a careful needs assessment, can establish the alliances to help them successfully establish an institutional repository.

A significant question is whether an IR is likely to serve the purpose most often proposed – to change scholarly communication. (Crow, 2002) Changing scholarly communication will require attention to intellectual property, tenure and promotion policies, and a host of inter-related issues that are beyond the control of libraries. This paper explores how libraries can position themselves to more effectively influence those debates by creating services for a range of campus denizens, fostering good will, and demonstrating that the library understands the needs of the entire campus community. In this way, the library may become a change agent on campus.

PART TWO: DEVELOPMENT AND DEPLOYMENT OF IRs

Institutional repositories are generally understood to have four primary characteristics. They are: institutionally defined (as opposed to discipline- or subject-focused); scholarly (containing the products of faculty, research staff, and students); cumulative and perpetual (the content will be preserved on a long-term basis); and open and interoperable (attentive to the Open Archives Initiative—Protocol for Metadata Harvesting). (Johnson, 2002)

Perhaps the most capacious definition of institutional repositories comes from Clifford Lynch of the Coalition for Networked Information. He writes:

> A university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution. While operational responsibility for these services may reasonably be situated in different organizational units at different universities, an effective institutional repository of necessity represents a collaboration among librarians, information technologists, archives and records mangers, faculty, and university administrators and policymakers. At any given point in time, an institutional repository will be supported by a set of information technologies, but a key part of the services that comprise an institutional repository is the management of technological changes, and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services. An institutional repository is not simply a fixed set of software and hardware. (Lynch, 2003)

This definition is widely used, as it not only encompasses the four characteristics identified by Johnson, above, but also underscores the collaboration required to sustain an IR. Lynch does not attempt to specify the content types or services that may be part of an IR, nor the technology that will run it. Most important for my purposes, he emphasizes the need for collaboration and an organizational commitment to long-term management.

This need for collaboration from across the university has been emphasized by OCLC on a more cautionary note: “Too few initiatives include all the stakeholders—faculty, library staff, IT staff and instructional designers—and there is no common view of what an institutional repository is, what it contains, and what its governance structure should be. (OCLC, 2003) Similarly, Roy Tennant, in a brief introduction to IRs in *Library*

Journal, addresses this point as a matter of economics. All the repositories he discusses began with support from their libraries. “How each of these … will be sustained over time may vary as much as the implementation models, but in all cases the long-term economic model is unclear.”(Tennant, 2002) Definitions of IRs that do not consider the need for sustained financial and administrative support cannot be considered adequate.

How does an institutional repository differ from a content management system? A briefing paper prepared by the Joint Information Systems Committee, (JISC) a UK group supporting information technology in higher education, explains that a repository is a *type* of content management system, but with a broader purpose than short-term access: “Typically a content management system holds resources for a particular course or departmental website. A repository can hold a comprehensive set of core assets that can then be used in a flexible way for different purposes, such as teaching… underpinning a website, or collating a university’s research outputs across a particular subject area or period of time.” (Joint Information Systems Committee (UK), 2005)

All of these definitions allow many possible services and uses for institutional repositories, as there will be varying needs and users. Additionally, they all emphasize the aspect of preservation. The vast majority of content being created today is digital. Institutional repositories are being proposed as a tool to select, organize, and preserve that content. All of these are traditional library tasks. But in the case of digital assets, the literature is clear that collaboration with partners beyond the library will be essential.

The Extent of IR Development

Despite the wide variety of possibilities, the literature and the first software developments assume that IRs will be primarily as a home for pre- and post-prints of scholarly articles. In 2003 research conducted for a UK consortium of publishers and libraries, Mark Ware found that the vast majority of the content was textual. (Ware, 2004 p. 27) However, this may be changing quickly. A 2005 survey conducted by the CNI indicates that current repositories are either accepting or planning to accept a great variety of formats. Textual material in IRs now includes theses and dissertations, preprints, conference proceedings, working papers, e-books, journals, newspapers, and university records. Non-text formats include data sets, power point slides, digital images, video, and audio, musical scores, maps, blueprints, software, web pages, and more.(Lynch & Lippincott, 2005) This enormous range of possible formats raises many issues, not least of which will be the challenge of preserving access to them. Beyond plans, however, what do we know about the extent of IR development?

The number of Institutional Repositories is difficult to determine. Virtually anyone can download the open-source software such as DSpace or Eprints, and there are probably hundreds of small repositories. These are not truly institutional, but rather storage spaces for a single department or research program. Lists of IRs have been developed for descriptive purposes, by SPARC (http://www.arl.org/sparc/repos/ir.html) and by individual researchers (Gibbons 2005). There is a registry of DSpace instances on the DSpace wiki at http://wiki.dspace.org/DSpaceInstances, but this is voluntary and includes many non-institutional repositories. Ware analyzed a list of some 45 IRs, but

was not able to determine an overall total number. (Ware, 2004) As already mentioned, a 2005 CNI survey of 124 research universities in the US found that 40% of respondents have some type of institutional repository operating, while 88% of those that do not are planning to implement one. While their sample is only a small portion of all US institutions, the authors believe that among the more than 250 research universities in this country, IRs will become a standard campus feature. (Lynch & Lippincott, 2005) Similar research in Europe and Australia estimated that the proportion of universities with an IR varies from 5% in Finland to 100% Germany, Norway, and the Netherlands. (van Westrienen & Lynch, 2005)

Even less is known about the number of objects in IRs, because definitions are so variable. Ware estimated that an average of 277 records in the 45 repositories he assessed at the beginning of 2003. (Ware, 2004) Lynch and Lippincott's attempts to indicate the size of the IRs they studied in 2005 were even less definitive—reported estimates ranged from hundreds of thousands of objects to a few dozen. They determined that "no standard way of counting the content exists at present… [because] no two institutions are counting the same things." (Lynch & Lippincott, 2005)

Further work to determine what constitutes an institutional repository, and to define nature of the objects being stored within it, is clearly necessary before effective metrics can be established, and is likely to be very difficult. For now, we can see that there is a considerable amount of activity at universities around the world, and the ambition to do more.

Advocacy for IRs

Supporters of institutional repositories have argued from several perspectives, in the recognition that establishing IRs will require what amounts to a cultural change for universities and their faculty. (Genoni, 2004) JISC's approach, similar to that of OCLC cited above, has been to emphasize the potential value of IRs for the several stakeholders. They argue that, for universities, IRs "help institutions to develop coherent and coordinated approaches to the capture, identification, storage and retrieval of their intellectual assets. . . [and thereby] enhance opportunities for efficient use of existing research, increase opportunities for improved learning experiences and encourage collaboration within and between different disciplines and groups." For faculty and students, JISC stresses the advantages of "free sharing of information, encouraging collaboration and the widespread communication of institutional education and research activity." (Joint Information Systems Committee (UK), 2005)

SPARC, an affiliate of the Association for Research Libraries, has emphasized not only the benefits to universities, but also the opportunity to change the publishing system. In their position paper, "The Case for Institutional Repositories," SPARC states that IRs "respond to two strategic issues facing academic institutions: 1) they provide a central component in reforming scholarly communication by stimulating innovation in a disaggregated publishing structure; and 2) they serve as tangible indicators of an institution's quality, thus increasing its visibility, prestige, and public value." (Crow, 2002) SPARC has also argued from an economic perspective, proposing that open-access publishing supported by author publication charges (which would be funded by research

grants) could shift the publishers' cost model from reliance on subscriptions. (Prosser, 2003)

An argument that many hope will convince faculty to self-archive relies on the evidence that open-access research is more frequently cited than non-open content. The first article to make this claim focused only on the computer science literature, but was nonetheless treated as valid in a broader context. (Lawrence, 2001) Further studies, however, have reached the same conclusion for other disciplines. Antelman found that in philosophy, political science, math, and electronic and electrical engineering, open access increased research impact, as measured by citation rates in the ISI Web of Science database. (Antelman, 2004) Davis and Fromerth, in an analysis of articles in math journals, concluded that articles deposited in the arXiv repository received 35% more citations on average than non-deposited articles. (Davis & Fromerth, 2005) As Antelman concludes, for libraries that aim to persuade authors to contribute to IRs, "Data showing that freely available articles in their discipline are more likely to be cited is powerful evidence of the value of repositories as well as other open-access channels." (Antelman, 2004)

Beyond the library and research communities, there is growing political support for open digital repositories (institutional and otherwise). Much of this arises from frustration that the public pays for scholarship more than once, by supporting the research and then buying the published work. This position was prominently articulated in a 2004 report published by the British House of Commons, *Scientific Publications: Free for all?* The

report recommended that, “all UK higher education institutions establish institutional repositories on which their published output can be stored and from which it can be read, free of charge, online. It also recommends that Research Councils and other Government funders mandate their funded researchers to deposit a copy of all of their articles in this way.” (Pinfield & James, 2003)

Such political support is not necessarily effective, however. In October 2005, the Research Council of the UK acted on the House of Commons recommendation, and began to require that its grantors deposit their research results in open-access archives, but the requirement is “subject to copyright and licensing arrangements.” (Research Councils UK, 2006) That “subject to” clause leaves open a very strong possibility that publishers can restrict work from open access repositories.

Similarly, the US National Institutes of Health established a Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research (Public Access Policy), effective May 2, 2005. Having been the object of considerable opposition from the publishing industry, this policy is even weaker than the RCUK’s, as it merely “requests and strongly encourages all investigators to make their NIH-funded peer-reviewed, author’s final manuscript available to other researchers and the public through the NIH National Library of Medicine’s (NLM) PubMed Central (PMC) immediately *after* the final date of journal publication.” (National Institutes of Health, 2005) A report to Congress on the first eight months of the policy indicates that it is so far a failure: “The

rate of submission to the NIHMS system… has been less than 4 percent of the total number of articles estimated to be eligible." (National Institutes of Health, 2006)

Clearly, such policies by research funders will need to *mandate* deposits if they are to have any considerable effect. Another alternative is for universities to exert some clout on their own. In the UK, the School of Electronics and Computer Science at the University of Southampton made self-archiving in their IR mandatory in January 2003, with the result that a high proportion of the eligible publications are now available. (Swan & Brown, 2005) Publishers are also responding to the open-access call in various ways. Oxford University Press, for example, has reported on experiments in shifting journals to "partial open access" (with some articles available free on the web after payment of voluntary author fees), "full open access," and permission for author self-archiving in institutional and subject repositories. (Richardson, 2005) In an attempt to provide comprehensive information on publisher self-archiving policies, the UK research group SHERPA maintains a searchable database of them at http://www.sherpa.ac.uk/romeo.php. They identify a sliding scale of author freedom: *green* publishers allow authors to archive pre-print (pre-refereed) and post-print (final draft, post-refereed) articles; *blue* publishers allow archiving of post-prints; and *white* publishers do not allow archiving at all. (Gold publishers, which are not listed in the SHERPA database, are fully open access.)

However, in a further complication, it appears that publisher policies have less impact than disciplinary practice. New research by Antelman found that authors publishing in

leading journals in six social science disciplines are self-archiving according to the norms of their respective disciplines rather than following self-archiving policies of publishers. Her research finds significant levels of self-archiving, as well as significant self-archiving of the publisher PDF version, in all the disciplines investigated. The variations were by discipline, not publisher policy. Publishers' self-archiving policies, she concludes, have no influence on author self-archiving practice. (Antelman, 2006)

Libraries that wish to open institutional repositories will need to monitor this rapidly-changing situation of research funders' mandates, publisher policies and disciplinary norms in order to advocate most effectively for their services. I address this issue in relation to specific stakeholders in Part Three.

Strategies for Attracting Content

One of the most prominent issues in the IR literature has been the difficulty of attracting content. Based as they are on the model of disciplinary repositories, IR content was originally expected to come from author self-archiving, but many scholars evidently do not see obvious benefits for depositing their work. It is becoming clear that authors' priorities, work habits, and disciplinary cultures may limit their initial interest in IRs. Hence a variety of strategies for attracting content are being developed.

Much of the research on this topic has come from the UK, where government funding has supported research on IR implementation. The DAEDALUS project at the University of Glasgow, for example, has addressed faculty concerns about whether they had the

right to deposit their work in an IR by having the library clear copyrights with the publishers. (M. Mackie, 2004b) As mentioned above, the SHERPA project at the University of Nottingham has responded to this concern more broadly, by producing a searchable database of publishers' copyright policies, so that authors and IR staff can determine rights without having to make time-consuming individual requests.

As they have implemented the DAEDALUS project, the Glasgow researchers report that "one of the major strategic shifts of the project is the decision to move away from encouraging academics to self-archive content, to the provision of a fully-mediated service." Accordingly, they have established three different repositories: for ePrints (peer-reviewed post-publication articles); for gray literature; and for electronic theses and dissertations. Each of these has required different strategies and services to gather content. Such advocacy is "an iterative activity that requires an investment of staff time." (Ashworth, Mackie, & Nixon, 2004)

The Australian librarian Paul Genoni writes from the library rather than the faculty perspective. He argues that managing IR content can draw from the familiar practices of library collection development, but with distinctions. Libraries have traditionally focused on purchasing finished materials from outside their institution for the use of those within it, but an IR is intended to develop material from within the institution for the use of those within and well beyond it. Consequently, "negotiations regarding content suitable for inclusion in institutional repositories will require librarians to become more familiar

with the full life cycle of scholarly research, communication, and publishing." (Genoni, 2004)

Research on the IR at the University of Rochester confirms both points: the necessity of understanding the research cycle so as to provide service to faculty as an inducement. There, a close examination of scholars' work practices revealed that building tools into the IR software to facilitate the writing process (including collaboration, versioning, etc.) could encourage faculty to incorporate the IR into the way they already work, which should result in more a more natural flow of content into the repository. (N. F. Foster & Gibbons, 2005). As a result, the library at Rochester is addressing the possibility of adding collaborative authoring tools to the IR, but these are complex. Such tools are under development elsewhere for a variety of purposes, and require attention to a large number of factors, including simultaneity, communication, workflow, security, intellectual property rights, and more. (Adler, Nash, & Noel, 2005) For the meantime, Rochester has created individual researcher pages, to facilitate the organization of content in their IR and enable authors to present a public face on the web.

In addition, the librarians at Rochester found that their initial descriptions of their IR did not strike a chord with faculty members. Terms such as *metadata*, *open-source software*, and even *institutional repository* met with disinterest or near-complete lack of understanding. The library now recruits faculty content by stressing the benefits from a faculty perspective: that the IR can make their work more easily accessible through web searches, preserve it safely without authors having to perform complex backups,

eliminate the need to save and send copies to colleagues, and allow authors to maintain ownership of it. (N. F. Foster & Gibbons, 2005)

Trust and Preservation

Beyond the problem of motivating individual deposits into the IR, there are broader issues encompassing their creators and potential users. The open-access literature reflects a debate over the short- and the long-term objectives, between Stephen Harnad's emphasis on filling IRs to support the development of new repositories and allow immediate access to research, with Peter Hirtle's concern for the long-term reliability of digital archives as central to their success. (S. Harnad, 2003; Hirtle, 2001) Pinfield and James point out that these are not mutually exclusive positions, but suggest moving forward on both fronts at once. They recognize that the long-term preservation of digital materials will require not only solving technical problems, such as how to preserve multi-media formats, but also the development of new organizational methods, including funding and decision-making processes about what to preserve for how long. (2003)

The twin problems of technology and management are found in other literature as well, linked to the issue of trust. The digital environment invites skepticism about the authenticity of objects. This is being addressed in research on the technologies of digital preservation, including digital signatures, persistent identifiers of objects, audit trails, and the creation of appropriate metadata. (Jantz & Giarlo, 2005)

More broadly, there is discussion of what constitutes a “trusted digital repository.” The Research Libraries Group and OCLC have produced a detailed report indicating the qualities of such an organization:

> A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. In this report the working group has articulated a framework of attributes and responsibilities for trusted, reliable, sustainable digital repositories capable of handling the range of materials held by large and small research institutions. The framework is broad enough to accommodate different situations, architectures, and institutional responsibilities while providing a basis for the expectations of a trusted repository. The critical component will be the ability to prove reliability and trustworthiness over time. (Research Libraries Group, 2002)

The attributes have mainly to do with sustainability and viability—technical, financial, and administrative. Trust is defined in terms of technical rather than social considerations. The responsibilities are in some ways familiar library concerns (scope of collection, preservation and other curatorial matters, the community to be served) but in the digital context also include somewhat distinct issues such as ownership of material (some of which may be digital copies of journals, for example), maintaining control of information in terms of the digital objects and their parts, and the support of best practices in the creation of digital resources. The report also addresses certification of trusted digital repositories, specifically focusing on models of individual, program, process, and data certification. (Research Libraries Group, 2002) The certification processes would be built on accepted standards such as those the Academy of Certified Archivists, the Historical Manuscripts Commission (UK), and the International Standards Organization. In a further development, the Center for Research Libraries was awarded a

grant in 2005 from the Mellon Foundation to develop the processes and activities required to audit and certify digital archives. (Center for Research Libraries, 2005)

Economics and Incentives

Much of the IR literature avoids the enormously significant issue of cost. Beyond the so far unknown expense of building and maintaining a digital repository lies the more general question: who is to pay for these services? The scholarly communications crisis arises out of economic paradoxes – faculty who give away their intellectual property to highly profitable commercial publishers, libraries unable to afford the breadth of content they would like, legislators indignant at the illogic of the system. The open-access movement is, not least, a search for new cost models for scholarly production and distribution. But most of the literature does not address costs specifically.

An exception is Brian Lavoie, an economist and researcher at OCLC, who has examined the economics of digital preservation from the perspective of incentives. Lavoie states:

> The fundamental economic issue associated with digital preservation concerns the incentives to preserve digital materials. The incentives to preserve can be characterized as perceived motivation sufficient to 1) induce a party to recognize a need to take action to secure the long-term viability of digital materials in which they are a stakeholder, and 2) induce a party to develop and implement technologies aimed at ensuring the long-term viability of digital materials. These incentives impact three key economic decision-makers in the digital preservation process:
> - Rights Holder: holds the intellectual property rights to the digital materials
> - Archive: provides services to ensure the long-term preservation and accessibility of digital materials
> - Beneficiary: benefits from the long-term retention of digital materials.
>
> (Lavoie, 2003)

Lavoie notes how incentives may vary depending on the organizational structure of any particular repository, and points out that different decision-makers may have quite different stakes in digital preservation. For example, “those who own digital materials are, in many cases, different from those who would benefit from their preservation: in other words, the Rights Holder is an entity distinct from the Beneficiary.” Furthermore, different levels of preservation may be desired by different stakeholders. All of these variations will influence the costs of preservation.

Lavoie’s analysis contains many insights for developers of institutional repositories. Libraries must seek a thorough understanding of the incentives of the IR stakeholders in order to motivate them effectively. The final section of my paper examines each stakeholder in term, in order to consider the path toward successful collaboration.

PART THREE: STAKEHOLDERS

What purposes will institutional repositories serve for different segments of the university community? If we consider an IR as a product, what jobs can it do for the university? Those outside the library may not recognize that the university’s digital assets are in need of curation and preservation. Librarians who wish to develop an institutional repository may have to persuade their university colleagues that the tasks of writing and disseminating research, preserving student records, creating and sharing curriculum materials, documenting administrative decisions, and so on, are at least in part problems of digital asset management. Each stakeholder – the library, archives, faculty, students, information technology managers, university press, and the university administration –

has distinct and overlapping needs in relation to digital material. In this section, I will investigate these stakeholders' digital information problems and the ways that an IR might – or might not – help to solve them.

The University Library

The fundamental role of a university library is to provide the intellectual resources to support the research and teaching needs of its faculty and staff. How can an institutional repository support this basic task?

Service to users: All libraries are challenged to meet their patrons' changing expectations. For an academic library to remain a vital and well-funded member of the campus community, it must provide resources in ways that its patrons would like to find them. In the context of the internet's vast and easily-reached content, the campus library is no longer the automatic source of information and knowledge. As discussed in Part One, the changing information culture calls for libraries to offer its services and collections in new ways. Library 2.0 is a vision of providing seamless access to a wide array of sources for a far-flung population of users. How can an institutional repository support those goals?

Collection-building: Libraries are in the business of building collections, of which IRs are potentially a new type. In addition to gathering community-created material that may have been published elsewhere, IRs may enable libraries to gather and organized unique

materials such as gray literature, datasets, multimedia works, and student scholarship, as well as digitized copies of analog material that the library already owns.

Such unique collections are gaining importance as a measure of library quality. As more users want material in digital format, libraries are recognizing that they cannot be first and foremost storehouses for bound volumes. Print is being transformed and moved off-site. Many special collections are being digitized to enable wider access. Regional storehouses for shared print collections are frequently mentioned as a solution to problems of limited space and tight budgets. The TRLN libraries, for example, are planning to build a joint storage facility, enabling the institutions to trim collections to a single copy of many journals and books.

As a result, one traditional measure of library quality – size of physical collections – may become less relevant. The Association for Research Libraries, long been the source of key ranking statistics, has established a task force on New Ways of Measuring Collections, among whose goals is "to develop a profile of the characteristics of a contemporary research library using qualitative methods that could serve to complement [quantitative] measures of library collections." (See http://www.arl.org/arl/ala06report.html#37) All of this suggests that unique holdings, including special collections, will be increasingly significant as a measure of library value. Along with rare books, manuscripts, and archival holdings, institutional repositories with their born-digital and digitized items could provide a new source of unique collections.

Preserving access to licensed content: A growing proportion of library holdings is licensed. In addition to its high cost, there is the problem of retaining access to this material once the license ceases. Libraries make efforts to negotiate licenses that promise "perpetual access" but many providers simply cannot guarantee this – nor is it in their interest to do so. The Digital Library Federation recently released a call for libraries to address this problem. Among the actions urged was that "libraries must invest in a qualified archiving solution. A library may itself operate a qualified archive... [or] research and academic libraries may collaborate."(Waters, Donald J., editor, 2005) There have also been calls for libraries to "demand archival deposit by publishers as a condition of licensing electronic journals."(A. Foster, 2006)

Several solutions are being developed. LOCKSS (Lots of Copies Keep Stuff Safe) uses open-source software to enable participating libraries to archive digital copies of serial publications (with the publishers' agreement) so that if the publisher is temporarily or permanently unable to deliver, there will be an alternate source. Portico, a similar publisher/library partnership, is an electronic archiving service to guarantee long-term access to serials. The methods and cost models of these two systems differ, but their purpose is fundamentally the same. Institutional repositories may also address this need, supporting these efforts or others to ensure reliable distributed digital preservation.

Promoting open access: Among the original justifications for building IRs is that they support open access. Consequently, major funders such as the NIH and the Research Councils of the UK consider institutional repositories one potential home for long-term

storage of research results. This may be accomplished by posting copyrighted material after a set embargo period, or by authors self-archiving their work (pre- or post-publication). For libraries, the advantages are not only the possibly lower cost of acquiring this material but also the possibility of preserving it.

Advantages and challenges: Building an institutional repository would allow libraries to address all these concerns—improved service, collection-building, preserving digital content, and supporting open access—but will require them to take on new tasks. Digital preservation, for example, requires an ongoing commitment to refresh or migrate materials into new formats, with unknown long-term costs. Similarly, the push for libraries to collect and disseminate unpublished material (pre-published articles, datasets, new open-access journals, and so on) may call for new roles such as peer review that are generally coordinated by publishers. Libraries will need to recognize that the development of an IR entails new roles, new budgets, and the development of job skills their current staffs may not hold. I will explore this issue more fully in my conclusions, in the context of other stakeholders' needs.

The University Archive

While archives are often administratively part of the library, it is worth considering their needs separately, as they are somewhat distinct. Their primary role is to acquire and preserve institutional records. As most such records (e.g. memos, emails, databases) are now born-digital, the campus archive is an obvious stakeholder in the development of an institutional repository. The university archive must expand its charge, or at least its level

of collaboration with others on campus, in order to responsibly manage the curation of these records.

Consciousness-raising: The collection of digital materials may require new levels of collaboration with other campus stakeholders who may not be considering the problems of long-term digital preservation. For example, the entity in charge of campus computing may be responsible for key administrative records of student registration, finance, and personnel. As Leon Stout reported on a case study of electronic records appraisal at Pennsylvania State University, “In general, there is little or no tradition of administrative computing centers and academic archives working together.” (Stout, 1995) That study found that although electronic records had been dutifully backed up to tape, only 1% of them were suitable for archiving, mostly because of inadequate documentation. The implication is that university archives must find ways to educate campus decision-makers about the very real danger of losing data if archiving is not made part of campus information planning. Such planning includes everything from purchasing systems that enable digital preservation to recognizing that electronic records management begins at the moment records are created.

Establish policies for the curation of digital material: A consequence of consciousness-raising is policy-making. The archivists and records managers cannot on their own determine what needs to be saved, and for how long. Furthermore, even a policy developed in careful consultation may be very hard to enforce, and different departments and units of the university may require different policies: the legal issues for

medical records will be considerably different from those of say, the humanities' departments. Each individual faculty and staff member will have to be trained in procedures for saving certain documents and providing adequate metadata to make the records useful to others. The more such tasks can be automated, the better, but there is no escaping the need for human engagement. Efforts to study this issue and create standards are underway, including "Managing the Digital Desktop," a National Historical Publications and Records Commission (NHPRC) grant-funded project to study computer file management practices in academic units and administrative offices, across the 16-campus UNC System and at Duke University (see http://ils.unc.edu/digitaldesktop/).

Advantages and challenges: For university archives, the creation of an IR is not the solution to the essentially social and cultural challenges of educating the administration, staff, and faculty about the need for digital preservation. But the repository can be a tool that demonstrates the archives' ability to provide reliable storage for digital records. Even more, the opportunity to build an IR enables new collaborations – the archivists can learn more the workflows of those who create and manage university records, and demonstrate their readiness to solve problems beyond their own immediate needs. Archivists and records managers bring their invaluable experience of appraisal and curation to the process, as well as their expertise in records management. In the process of establishing an IR, the archives can assert their value to the university as a whole.

Faculty

The IR vision assumed the participation of faculty by self-archiving their work, but the reality is that many IR administrators have struggled to attract content. It seems likely that faculty could benefit from a convenient long-term storage system that provides open access, but perhaps they do not recognize these as needs. What does do faculty members need to support their work, and how can an IR help them?

A general sense of faculty awareness of open access and self-archiving may provide some context. Researchers in the UK (funded by the Joint Information Systems Committee) provide some valuable data. (All data from Swan & Brown, 2005) They surveyed 1,296 authors from around the world and across the disciplines, and with varying levels of knowledge and experience of open access. They found that 49% of their respondents have self-archived at least one article, 27% by simply putting a copy on a website, with 20% in institutional and 12% in subject-based repositories, although use of IRs and disciplinary repositories is growing fast. Of the authors who have not yet self-archived, 71% are unaware of the option (amounting to 36% of the total author population). Concerns about self-archiving include the perceived difficulty of doing so (which drop off dramatically after the first act of depositing) and worries about copyright infringement. Interestingly, however, many authors seem not to do much about the latter concern. Among self-archiving authors, 17% seek permission if they believe it is needed. Only 10% of authors know about the SHERPA list of publisher policies. But where it is not known if permission is required, authors are not seeking it and are self-archiving without it. (This jibes with Antelman's finding, previously noted, that publisher policies have little impact on self-archiving practice.) Perhaps most interesting is the finding that

the vast majority of authors (81%) would willingly comply with a mandate from their employer or research funder to deposit copies of their articles in an institutional or subject-based repository.

These findings, presented in considerable detail and with breakdowns by subject area, are a useful starting place for understanding faculty knowledge and behavior. Universities differ, and faculty needs will, too. Libraries will have to assess their campus communities, identify services that can be offered through the IR, and locate likely early adopters from whom they can learn more. Following are specific areas of likely faculty concern.

Create, revise, and store research: In the course of conducting research, faculty members create digital files of many kinds, from drafts and data to final versions. IRs have been designed mainly to provide access to the finished work, with the goal of enabling wider access. However, research at the University of Rochester "revealed some of the reasons why current IR systems are more useful to faculty in theory than in practice."(N. F. Foster & Gibbons, 2005) As noted above (pages 24-25), this study found that faculty wanted more help with authoring than with storage – keeping track of versions, enabling access to their collaborators from different locations, and so on. Rochester is modifying its repository to make it more natural for faculty authors to use it.

Share research: All researchers want to share their results, whether with a small group of their closest collaborators or with a wider audience. However, this may not mean that

scholars know much about open access. Swan's study found that the main reasons why authors have not published in an open-access journal are that they do not know any in their field, or cannot identify a suitable one. (Swan & Brown, 2005)

Another barrier to self-archiving may be disciplinary culture. For example, medical research is expected to be rigorously peer-reviewed before being made public, because of potential risks to human life. Humanities fields in which books are the publishing standard generally do not expect or support quick access to results. Understanding such differences may help libraries customize their approaches to different campus communities.

Others argue that faculty are focused on their own work more than on the wider benefits that open access may promote: "many academics feel that assuring access to research is not central to their work. They are engaged in attracting funding, completing research and publishing the results. Finding their publications to build further research proposals is the problem of subsequent researchers. Most researchers want to publish their findings in the highest prestige journal possible, essentially … making a leap of faith that their contributions will reach a large audience and make a favorable impression on those holding the purse strings at institutions and grant making organizations. Even among scholars familiar with access issues, few feel that impacting the publishing paradigm is an important or attainable personal goal." (Corby, 2006)

However widespread these beliefs may be, research has consistently shown that material available in open access journals and repositories is cited more widely than non-

OA research. These results should be strategically promoted by librarians to faculty, as evidence that the institutional repository offers potential to increase their impact and influence. Some faculty will respond to self-interest, while others may be persuaded by arguments about the public good.

Support tenure and promotion: One reason for sharing results is the need to earn tenure and promotion, a crucial part of most faculty members' professional and economic lives. Published scholarship in journals or book series that the university considers appropriate and of good quality is a key requirement. This process depends upon peer review, generally conducted by publishers. It may seem that IRs do not support this need or even mitigate against it by co-opting the audience for the research. However, the decade-long existence of subject repositories suggests that open archiving does not present a danger to tenure. Moreover, depositing papers in open access repositories often leads to increased citations, a factor in tenure reviews.

Futhermore, IRs do not necessarily preclude peer review. Much content in the disciplinary repositories is posted previous to publication. And peer review can be conducted on IR contents. (S. Harnad, 1995) Some IRs have also created distinctions between content that is peer-reviewed and that is not. For example, the California Digital Library's eScholarship repository offers a service by which University of California units can publish peer-reviewed journals and series. These are clearly marked, and a description of each unit's peer review process accompanies each paper. (See http://repositories.cdlib.org/escholarship/peer_review_list.html). This service offers a

powerful demonstration that new methods of scholarly communication can be supported in the institutional repository.

Another way the IR may support promotion processes is as a tool: curriculum vitae can be linked to IR contents, enabling easy access for those who must review faculty work, and generally showcasing the work of individual scholars. The NCSU libraries, for example, have for many years provided faculty members with a database in which to store citations to their published works. This database is being used as a platform from which to move content in the campus IR. Such services should help to promote the IR.

Create, store, update and share curricula: Although this idea has not been much addressed in the literature, an IR could support the creation and sharing of curriculum material. It is a common practice for faculty to produce course web pages or use courseware systems like Blackboard and Sakai. An IR could be a source of content needed for teaching purposes, available by easy linking. Blackboard, in fact, has begun collaborating with BE Press, a producer of software for institutional repositories and electronic journal publishing, to create ResearchNow, a database of research articles that can be searched from within Blackboard. Many university libraries already provide links from course pages to licensed content or e-reserves. This is a service that faculty and students may have come to expect. The IR should be built with this function in mind.

Advantages and challenges: Constructing an IR that will meet faculty needs requires the willingness to understand their workflow and research methods. A library that can

collaborate effectively with faculty will build a powerful advocate for future library funding. IRs offer great – albeit perhaps so far untapped – potential for such collaboration. One place to begin may be the already-established lines of communication from the library to the faculty (such as subject bibliographers or other library liaisons). Another option would be to identify scholars who, out of disciplinary or other indication, may become early adopters.

One researcher posits that there may be a misunderstanding of what an institutional repository is:

> [T]he whole idea of self-archiving in institutional archives is based upon false assumptions about the behaviour of academic authors. Academics publish and the problem with the concept of an archive is that it is generally perceived as a mode of preservation, not a mode of publishing. Archiving also depends on the voluntary depositing of already-published, or about to be published, material, and some strategy is needed to ensure that academics collaborate. (Wilson, 2006)

How can libraries address these challenges? In part by educating faculty on the individual and social benefits of open access, in language most likely to connect with faculty goals and wishes. Libraries could rename the “repository” with a term that conveys the idea of distributing, rather than merely storing, digital content. Fundamentally, the IR should not only be described but should function as a useful tool for the faculty to use for the work that they are already engaged in.

The Students

Most students arrive on university campuses as experienced users and creators of digital content. According to a 2004 study from the Pew Internet and American Life project, the most enthusiastic creators of content on the internet have an average age of 25. (Lenhart, Horrigan, & Fallows, 2004) They write blogs, post photographs, and share video and audio files. The use of social networking sites is practically universal: research at UNC has found that 94% of the freshman class had a Facebook page by the end of their first semester. (Stutzman, 2006)

In relation to education, another Pew study found in 2001 that 71% of online teens reported using the internet as their major source for research. Nearly three-quarters (73%) of college students say they use the Internet more than the library, while only 9% said they use the library more than the Internet for information searching. (Lenhart, Simon, & Graziano, 2001) While this research does not assess how many students use the internet to find library materials, it indicates that online materials are most likely to be discovered by students.

Libraries are aware of the need to create systems that allow students to find and use online the material that supports their research and learning. The work done by libraries on metasearch, bibliographic instruction in the use of databases, and OPAC usability are all indications of efforts to serve students more effectively. Institutional repositories are another method with which to engage the student population. What are students' specific needs in relation to digital assets?

Find and use curricula: In the mirror image of the faculty need to create and enable access to materials for coursework, students must locate and use these items. The IR as a source of content for class material can meet this requirement. In addition, IRs may serve as a space for student-created curriculum materials and group assignments.

Construct and maintain resumes and CVs: Similarly to faculty, students need to document their intellectual and work experience. An IR could provide a home for digital objects that students wish to share with prospective employers or other educators.

While many of these materials may not be relevant for long-term storage, an IR could provide the possibility of spaces where students may create and store digital objects, whether for class use or other creative purposes. Time limits may be placed on some of this content (a topic addressed further below.)

Preserve major research projects: Students may be the population on campus most likely to contribute material to an institutional repository, because they can be required to do so. Theses, dissertations, and other work that is required for completion of degrees can be mandated for storage in the institutional repository. This is merely a shift in requirements for formatting and submitting copies of such works, with the advantage of making them much more widely available. In fact, storage of theses was among the first uses of institutional repositories, starting with Virginia Polytechnic University's Electronic Theses and Dissertations Project, which began development in 1987 and has now evolved into the Networked Digital Library of These and Dissertations,

http://www.ndltd.org/. This paper, in fact, must be deposited into the digital repository for the School of Information and Library Science at the University of North Carolina at Chapel Hill.

Advantages and challenges: Student-made content is a useful starting place for IRs, because deposit of it can be mandated. Institutions that are already accepting electronic theses and dissertation will have a ready supply of available material. Such work can, in effect, seed the repository. Other than mandated material, however, attracting student content will, as with faculty, depend on the ranges of services offered.

Student work also suggests some possible adjustments in the notion of an IR as a permanent home for preserved digital content. Not all student content will be available, as some of it may be on restricted sites (Facebook, for example, runs on proprietary software) and other student projects are simply not appropriate for long-term preservation. Such issues create the opportunity to establish variations in time limits for IR content, and the feasibility of linking outside the IR. Additionally, the IR can also be a source of education about copyright and intellectual property issues. Students and universities are already grappling with policies related to downloading music and other forms of filesharing. Establishing an IR may open a wider conversation about the available range of licenses (such as Creative Commons), the choices that content creators face in permitting access to their work, and fair uses of copyrighted material.

Students may also place special demands on an IR, as they are likely to use a wide variety of media and formats for their work. Much of the current IR software, such as DSpace and Eprints, is designed to accept mostly text. Repositories that wish to support multiple formats (whether for long-term preservation or only during a students' years on campus) may learn a great deal from student work.

Information Technology Services

Contemporary University IT departments have a very large job, providing computing support for the complex enterprises of research, teaching, and administration, with services for everything from student registration, human resources, and finance to laboratory and classroom technology. IT staff maintain the applications that campus users need—email, telecommunications, calendars and directories—as well as campus networks and servers. They are also likely to be responsible for the data that supports all activities, including backups and disaster recovery, and the creation of policies and planning for all of the above. Their work requires a high level of service and competence, under circumstances of constant change.

Despite much talk of potential synergy between libraries and campus IT departments in the 1980s and 90s, when "electronic libraries" began to be envisioned, most large universities have kept these departments administratively separate, with IT generally headed by a Chief Information Officer and the library reporting to an administrator for Academic Affairs. (Bolin, 2005) Hence the priorities and plans for the two departments may be created and implemented without consideration of the other's needs and

expectations, with consequences for potential collaboration. How might the range of IT tasks benefit from an IR?

Supporting research, teaching, and learning: The IT division is generally responsible for the hardware, software, system administration, and data storage that is required for the research and teaching. With a focus on the technical aspects of these chores, IT staff may not distinguish between items that need storage for varying amounts of time. The library and archives, on the other hand, as the campus's key content selectors, are interested not only in making that content available but also in preserving some of it indefinitely.

IT staff are generally involved with supporting the courseware systems. As mentioned in regard to faculty, courseware systems enable linking to e-reserves and licensed content. This capability could be extended to content in the IR. This service is clearly within the missions of both the library and the IT service areas.

Preserving digital records: University IT departments are generally focused on the present, responding to ever-increasing needs for data storage, the daily challenges of back-up, and planning for disaster readiness. As mentioned on university archives, IR proponents may need to raise IT consciousness about the challenges of long-term digital preservation. IT administrators may not be aware of how easily their data, however safely stored, can be lost if its organization is not effectively documented and its content regularly migrated. IT's mission of backups and archives' mission of preservation mesh

naturally. This partnership has not been exploited to the extent possible in most institutions.

While IT departments take on the responsibility of storing research data, university records, and other digital content, they may not consider the legal requirements to store such data indefinitely. Increasingly, funders such as the National Institutes of Health, the Wellcome Trust, and the Research Council of the United Kingdom are creating or contemplating policies that the results and data of research they support must be made available in open access repositories. Most of that research data now resides on individual faculty computers and departmental servers. IT departments will need to collaborate with those creating IRs to enable the orderly movement of content into the repository.

Establish policies for and create systems that enable digital preservation: One of the key problems of preserving digital content is the need to do it in the moment of creation, by providing adequate descriptive, structural, and administrative metadata. It is generally expensive and time-consuming to add such metadata later – and it may not even be possible. University IT departments need to consider which records need to be saved, and for how long; and they need to create policies and workflows that enable this. IT managers may need to consult archivists and records managers as they consider the choice of software and hardware systems that store data, as well to ensure that data can be extracted from them and preserved in useful ways. An IR can support a system by which policy documents, course catalogs, and other material that is regularly revised can

be stored for historical reference. The development of an institutional repository can instigate such cooperation. There are benefits for campus IT leaders:

> The CIO has an obligation to ensure timely and effective campus response to technology issues. Academic digital assets—maintained, as they usually are, in a decentralized way—should be of deep concern to the CIO. Implementing an IR is a way to begin to explore the role of central IT in aiding digital preservation and in opening up a campus discussion on these topics. (Goodyear & Fyffe, 2006)

The University Press

The last few decades' shifts in scholarly communications practices have affected university publishers dramatically. How and whether university publishers will move into digital publishing is an open question. Not all universities support a university press—there are 129 members of the Association of American University Publishers, and more than 250 research universities in the U.S.—but these presses are key players in scholarly communications. This section will address their particular needs and concerns in relation to digital material, and their possible benefits from an IR.

Declining support: University publishing began in the U.S. in the late nineteenth century, as a means for universities to gather and distribute their faculties' research in both book and journal form on a non-profit basis. The core market for such work was university libraries, first by means of exchange and eventually for purchase. This system thrived in the post-World War II era, as government funding for education was plentiful. With science and technology research on the rise, libraries began to shift their acquisitions budgets into serials, the primary means of publishing in those fields. Funding for monographic scholarship declined, at the same time that serial prices began rising

steeply. University presses, particularly those that depended largely on monographic sales in the humanities and social sciences, took a powerful hit in a key market sector. By the 1990s, sales of scholarly monographs to libraries had dropped from an average of 1000 per title to around 300. (Givler, 2002) At the same time, university subsidies to presses declined or were cut altogether, as university administrators looked to their publishing units to cover their own costs. (Regier, 2003) Presses have, in response, sought to support their finances by publishing works that have a larger-than-specialist market (books for general readers and for the classroom) and by raising endowment funds. (Givler, 2002) The close link between university libraries and university presses began to erode as the publishers were forced to seek markets elsewhere.

Need for technology support: As will be clear from this brief history, most university publishers operate on very tight budgets. Publishing is traditionally a small-profit business, and specialized scholarly publishing addresses particularly small markets. As a result, university presses maintain tight control over their costs, and are generally unable to invest in innovative technology or hire staff members solely devoted to experimental digital projects. Most of the digital-publishing efforts in the US university press world have been supported by grants (primarily from the Mellon Foundation).

Need for sustainability: As university publishers face steadily decreasing monograph sales and competition with for-profit journal publishers, they will need to find ways to support their missions over the long term. One researcher has described it this way:

> Scholarly book publishers are caught between their scholarly missions and business pressures and opportunities. The business pressures include

> library budgets increasingly devoted to expensive journals rather than books, margin pressure from a few increasingly dominant retail booksellers, and more efficient used-book markets thanks to the Internet. (Nitterhouse, 2005)

In response to these pressures, publishers have sought ways to trim their costs and take advantage of new technologies. For example, the University of Chicago Press (with $2.75 million in Mellon Foundation support) created BiblioVault, which charges fees to archive digital book files for its 36 member presses (http://www.bibliovault.org/). This was envisioned in the original grant proposal as a cost-effective way for publishers to manage digital files for print-on-demand services. (BiblioVault, 2000) At this point it seems to operate in this way – there appear to be no plans to make use of the digital files for online distribution. However, it is possible that such a service could be developed, should publishers be willing to risk it, and to solve such practical issues as licensing procedures for small bits of digital content. But new experiments in distributing ebooks are underway by entities from commercial publishers, the Open Content Alliance, Google, and Amazon. The university press community may ultimately need to respond to these ventures, and institutional repositories may be able to support them.

There are already a few library/university press collaborations underway. The California Digital Library eScholarship repository, has partnered with the University of California Press, to the benefit of both. The CDL provides storage space and technology to deliver UCP e-books, while the Press conducts peer review offers its imprint to some eScholarship journals and series. The e-books (mostly backlist, not new titles) are available online as eScholarship editions, while the journals and other series are in the

eScholarship repository. This partnership is intended to support new ways of distributing scholarship, although there are some constraints: the content of most of the 1,500 books in the eScholarship program is accessible online only to members of the UC community, not the wider web audience.

A similar project is being undertaken at Pennsylvania State University, in the Office of Digital and Scholarly Publishing, a press/library collaboration. Among other projects, they have created online and print-on-demand access to the press's older titles. (See http://www.libraries.psu.edu/digital/scholarlycomm/) Another example is at Georgetown University, where a new partnership between Georgetown University Press and Digital Georgetown will offer free access to scholarly books in a digital format to students, faculty and staff at Georgetown University. (See http://www.library.georgetown.edu/dept/admin/librarynews/1-25-06news.htm#news)

Most of these projects do not create open access to published materials. Doing so for older in-print titles would allow easy discovery of the books at minimal marketing expense, and may create a small print-on-demand market. Some university presses are collaborating in the Google Book Search program, by which books can be discovered and browsed online, with the hope that sales will increase. Another option would be to post out-of-print books in IRs. Such ventures allow a considerable amount of orphaned material (that is, under copyright but not being actively distributed in print) to remain available. This would cost the university presses little, but may earn them considerable goodwill, an important consideration in their quest for long-term sustainability.

Advantages and challenges: As suggested by these partnerships, university presses have skills to contribute to institutional repositories. However, there has been little movement in this direction so far. This seems to be the result of very different approaches by libraries and university presses to open access. Most publishers, including university publishers, are not in the business of making their content freely available. Working within very tight constraints, university presses seek to find every new income stream they can. There is great concern that the already-small markets for scholarly content will vanish if it is made freely available. Providing IRs with free content may seem counter-productive from this perspective. Certainly there has been considerable (although not unanimous) opposition in the university press community to the Google Book Search project – the Association for American University Presses has joined the suit against it. The largest challenge for libraries in collaborating with university presses may be cultural and economic. Librarians need to understand the publishers' perspective on control of content in order to work effectively with them.

In addition, librarians can educate themselves about what publishers are best at. Many discussions of the potential for digital publishing seem to assume that publishers, having obtained the copyright to high-quality content at little or no cost, do very little beyond formatting and distributing it in published form. Granted, the main role of scholarly publishers is acknowledged to be one of quality control, supported by their practice of peer review (on which more below.) However, publishers can also provide other valuable skills, including the establishment of selection criteria; experience working with scholars as they shape their work; careful editing; and marketing or promotion to diverse

audiences. All of these require experienced staff. Acquiring and publishing content, in that sense, is a considerable expense. Not all publishers undertake all these tasks at an equally high level, and the processes and cost models differ for journals and books. But serious scholarly publishers add considerable value to their products. And most university presses and scholarly societies charge only enough to cover their costs.

Peer review deserves special mention here, because the entire system of scholarly communications relies on it. The “publish or perish” requirements for tenure and promotion assume the existence of an impartial review process by which scholarly work is determined to be of publishable quality. Early concerns in the academic community assumed that scholarship distributed online would not be rigorously reviewed. Institutional repositories, as a home for self-archived work, may be the object of similar concerns about quality control. However, commentary on self-posted work is, of course, a form of peer review. Discussions of these issues should acknowledge a spectrum, with one end perhaps being self-publishing on personal websites and the other formal publishing by well-regarded organizations after rigorous peer review. Content in IRs may fall along many parts of this spectrum: peer-reviewed publishers’ versions of material; peer-reviewed by the institution hosting the IR itself; or self-posted, awaiting commentary, and eventually resulting in publication. Those supporting IRs can make clear that open access and peer review are two separate issues, and, where appropriate, distinguish the review level of content in their repositories.

University publishers can be enlisted as partners in an institutional repository, particularly if there is a wish to establish different levels of quality control within it. Publishers spend a considerable amount of time and energy building and maintaining their lists – that is, selecting the works they publish – precisely because their reputations rely on those tasks. The large role granted to publishers' quality control in the tenure and promotion process makes explicit the value placed upon publishers' reputations in the system of scholarly communications. Libraries that wish to engage their publishing colleagues need to acknowledge and value what publishers bring to the table.

Similarly, university publishers may want to acknowledge their libraries' influence on campus, and consider where their own long-term interests lie. The book publishers in particular now devote enormous energy to marketing and selling their books in commercial venues. For much scholarly content, this effort and expense does not significantly increase sales. Those publishers committed to the availability of excellent scholarship cannot expect to find reliable allies among bookstore chains. Collaborating with colleagues on their campuses may create good will and protect their ability to fulfill their mission as publishers.

Institutional repositories can offer benefits to both libraries and the publishers, provided that each party recognizes the others' priorities and concerns.

The University Administration

The primary ally for a library is of course the university administration, from whom the resources flow. What are some of the specific tasks that a university administration needs to accomplish and which an IR can support?

Support research, teaching, and learning: All institutions of higher learning must support their faculty and students. Clifford Lynch has pointed out in a recent interview that, “there has been huge investment over the last few years in infrastructure to support teaching and learning: learning management systems; smart classrooms of various types; all kinds of new learning spaces and, particularly, informal learning spaces.” (Hawkins, 2006) Such investment indicates a university priority and raises the question of how libraries should respond. As indicated earlier, an institutional repository can provide valuable support for courseware systems. Beyond that, a successful repository—one that contains a considerable quantity of high-quality research—can not only provide evidence of what the faculty and students are producing, but also enable such products to be used for further scholarship.

There are many possibilities for IR development here. Universities may offer space in institutional repositories for heterogeneous collections of data, for example, making them available to scholars and students everywhere. One small example of such a collection is “Inside Wood” at the North Carolina State University Libraries, where a scholar in wood sciences has worked with the library to digitize slides representing this slices of different species of trees, and make them available through a web-accessed database. These images are essential for the discipline, enabling identification and comparison, and

supporting teaching. Scholars around the world (in this relatively small field) have contributed their own slides to the project, making it ever more valuable. The challenges for such shared collections may be in preserving a variety of data and sorting out ownership issues for shared resources. But the benefits are also great, for scholars well beyond the home institution.

Retention of faculty: As James Moeser, the Chancellor of UNC at Chapel Hill, has stated, “A great university starts and ends with a great faculty. Thus, the number one priority for this University remains attracting and retaining the finest faculty in the world.” (Moeser, 2005) Derek Bok, during his tenure as the president of Harvard University, similarly made the hiring and retention of faculty his top goal. It follows that serving the faculty effectively, in the ways described above is a crucial way for the campus library to gain the good will of the university administrators. Libraries seeking support for IRs can make clear how the faculty has benefited, and recruit faculty to help them make the case.

Maintain control of university digital assets: Government agencies and the foundations that fund research generally require that data and research results from their funding be preserved and shared for further study. In addition to the policies of the NIH and RCUK discussed previously, the National Science Foundation grantees are expected to share “the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.” This goes beyond publication of results and analysis to the “responsibility that investigators and

organizations have as members of the scientific and engineering community, to make results, data and collections available to other researchers." (National Science Foundation, 2005)

Universities, therefore, have an obligation and a strong incentive (given the amount of research funds available from agencies such as NSF and NIH) to preserve research data. At any given university, identifying the major source of research funds may suggest what the institutional repository should target as a high priority. For example, at the University of North Carolina, the vast majority of research funds come from bio-medical research. At the same time, the greatest institutional risk rests with the loss of the data produced in that research. (Jones, 2006) Much of that data is scattered on individual scholars' computers and department servers. Advocates of IR development should articulate long-term preservation of such data as a university priority, and a service which the institutional repository may undertake.

Advantages and challenges: Persuading the administration to support the development of an institutional repository with financial and human resources is absolutely crucial, and will require the simultaneous creation of alliances and support from the other campus stakeholders. One great challenge in asking for support is that the long-term costs of an IR are impossible to predict. The expenses involved with digital preservation, for example, are unknown, but likely to be high, as content must be migrated or otherwise managed every three to ten years, a far higher frequency than with analog materials. In addition, the range of services offered as part of the IR will vary

according to campus needs and resources, but will surely require not only investment in technology and administration, but also the development of new employee capabilities in the library and elsewhere. Studies that can begin to assess these costs are sorely needed. One model may be Montgomery's D-Lib essay that examines how electronic publishing changed libraries' operational costs because of shifts in staffing, resources, materials, space and equipment. (Montgomery, 2000) Despite these unknowns, it can certainly be argued that given the near-complete reliance of contemporary communication on digital media, the costs of ignoring digital assets are likely to be far higher than managing them. No university can afford to risk losing the work of its community.

Clifford Lynch of the Coalition for Networked Resources has discussed the issue of repositories and data management in terms of university investment:

> CNI has spent much time on the question of what will happen to libraries and to campus public commons as we move into an increasingly digital world. In the late 1990s, a lot of institutions… invested heavily in an almost complete overhaul of administrative systems, partially driven by the Y2K threat. This proved to be a massive, and massively expensive, effort that I think took longer and cost more than most institutions expected—one that higher education is just now recovering from. The opportunity costs here were huge. Now, though, we're seeing a renaissance of interest in research computing. The economic, technological, operational, and policy trade-offs around centralization of computational cycles and storage have changed. Networking needs in research have also grown. Most important from CNI's point of view, however, is that research practices have changed quite a bit over the years. Research infrastructure is no longer about just cycles and networks. Today it is also about very-large-scale data management and about facilitating collaborations. (Hawkins, 2006)

It will be incumbent upon those who wish to obtain university support for an institutional repository to recognize the priorities the university has indicated with its

investments and, where possible, point out how the IR may support these. Only with significant financial support can institutional repository can become a crucial part of the structure of campus computing infrastructure, facilitating research, keeping faculty engaged, and providing campus-wide services in digital preservation.

CONCLUSIONS

This paper suggests that institutional repositories, created by librarians as one solution to the problems of restricted access to and lack of control over content, are more likely to succeed if they work not only for the library but also for the stakeholders across campus that have an interest in digital content. A careful needs assessment, going beyond the interests of the library, is an important element in building an IR. Developing an IR in this fashion should result in benefits for all these parties.

The stakeholders' needs and potential benefits are complex and overlapping. For example, students and faculty could use many of the same services, and may be motivated by similar incentives. Their need to create and share work is fundamental, and well-established mechanisms already exist for these processes, from server space on the campus network to conventional publishing outlets and online social networks. The IR will remain an afterthought unless it meets these needs in ways that are incorporated into the natural flow of research and authoring. The University of Rochester research cited above is an excellent example of understanding what the faculty knows about IRs and how they might use their services. Further work is reportedly underway on that campus to study students' needs. The benefits to campus authors could be many. One, resulting

from the education about intellectual property inherent in depositing work in an IR, is an enhanced ability to negotiate and control the availability of their work. Others may flow from the services that are established: efficiency in sharing results, wider access to research, stability of data.

There would also be overlapping benefits for the university archives and the IT department, most significantly an enhanced ability to work together in their shared mission of storing and preserving the university's digital assets. The technical knowledge from IT could effectively complement the selection and curation skills of libraries and archivists to address the burgeoning amount of born-digital files. This is a problem that grows more urgent with every day that passes, and the IR can reasonably be offered as part of the solution.

The university press may benefit from having an on-campus partner to assist them with digital storage, and furthermore, may benefit from the libraries' understanding of campus information needs. There are many possibilities for cooperation in new ways of creating and distributing digital scholarship, from which the university press could benefit by sharing the risks with their campus library. The library, in turn, could benefit by working with the local experts in the creation, shaping, and assessment of scholarship as it emerges from idea to finished product.

The university administration would profit from all of the above, by supporting a structure that reinforces the campus's efficiency, productivity, and intellectual reputation.

Providing better service to all these stakeholders is in the university's best interest. And maintaining control of digital records, research data, and results is a critical task in sustaining the university's ability to attract not only excellent faculty and students but also research funds.

And finally, of course, libraries could benefit in many ways. They would acquire content over which they can exert more control, enabling better delivery of it to their patrons. By collaborating effectively with these stakeholders, the library would also reinforce its position on campus as a vital partner in the university's overall mission. The library could establish itself as an innovator and change agent, entering fully into the scholarly communications process rather than simply responding to its shifts and trends.

I have posited many possible services for the IR, including authoring support, archiving a variety of formats, creating curation policies for different kinds of university documents and other data—as well as the ongoing IR administration and preservation tasks. No library could establish all of these from the outset, and many libraries may determine that other tasks are more appropriate. The idea explored in this paper is that a thorough needs assessment is the first step to determining what services the IR might offer to support different stakeholders.

Beyond this vision—needs assessments that lead to effective collaboration with benefits for all—there remain questions. As different IR services are introduced, it will be

important to evaluate their effectiveness. Likewise, developing a method to count objects in IRs would enable some broad-scale comparisons and cost estimates.

The underlying question is what purposes IRs should serve. Does the original vision, that they will create new methods of scholarly communication, seem likely? Perhaps it is best to consider this question in stages.

In the short term, IRs are unlikely to alter such a complex system. Changing scholarly communication will require shifts in intellectual property law and publishing—beyond the ability of universities to control. More possible would be changes in the ways that universities handle tenure and promotion. For example, it would help to find ways to evaluate and reward non-traditional forms of scholarship, particularly digital scholarship and online teaching. The challenge would be that no university can radically change the tenure promotion system unless many others universities begin to do the same. (Pomerantz & Blouin, 2005) For libraries, the problem is even greater, given that they may have relatively little influence on such issues even within their home institutions.

What libraries can do now is to begin building a set of high-quality services for digital content, to demonstrate that the university library can be a major player in a new scholarly communications regime. They will then have a product—the IR itself—and a set of campus partners who can vouch for the library's central role. For any library, such local allies are critical for ensuring adequate financial support.

For the longer term, we can hope that such services will create changes in the scholarly communication system. A well-developed IR could enable its library to address some of the problems—the increasing costs, the loss of control over their content, the demand for more digital material—that have suggested this possible solution. This paper has argued that in order to begin that task, a library will not only have to articulate its own needs and challenges, but also to serve those of its campus partners.

References

Adler, A., Nash, J. C., & Noel, S. (2005). Evaluating and implementing a collaborative office document system. [Electronic version]. *Interacting with Computers, In Press, Corrected Proof*

Antelman, K. (2006). Self-archiving practice and the influence of publisher policies in the social sciences. *Learned Publishing, 19*(2), 85-95.

Antelman, K. (2004). Do open-access articles have a greater research impact? [Electronic version]. *College & Research Libraries, 65*(5), 372-382. Retrieved 3/17/2006, from Library Literature and Information Science database.

Ashworth, S., Mackie, M., & Nixon, W. J. (2004). The DAEDALUS project, developing institutional repositories at Glasgow University: The story so far. *Library Review*, *53*(4), 259-264.

BiblioVault. (2000). *A Vision for Innovation in Scholarly Book Publishing: The Chicago Digital Distribution Center and BiblioVault.* Retrieved 4/10/2006 from http://www.bibliovault.org/vision.pdf

Blummer, B. (2005). Session report: ALA midwinter 2005, ACRL current topics discussion group, building an institutional repository. *Cites and Insights: Libraries, Policy, Technology, Media, 5*(3), 21--24.

Bolin, M. K. (2005). The library and the computer center: Organizational patterns at land grant universities. [Electronic version]. *Journal of Academic Librarianship, 31*(1), 3-11. Retrieved 2/4/2006, from LISA database.

Center for Research Libraries. (2005). *Auditing and certification of digital archives.* Retrieved 4/7/2006 from http://www.crl.edu/content.asp?l1=13&l2=58&l3=142

Christensen, C. M., Cook, S., & Hall, T. (2005). Marketing malpractice. [Electronic version]. *Harvard Business Review, 83*(12), 74-83. Retrieved 1/26/2006, from Business Search Premier database.

Corby, K. (2006). *Review of "the access principle: The case for open access to research and scholarship," by John Willinsky.* Retrieved April 1, 2006 from http://blog.lib.umn.edu/scholcom/accessdenied/042159.html

Crow, R. (2002). *The case for institutional repositories: A SPARC position paper* No. http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf). Washington, DC: Scholarly Publishing & Academic Resources Coalition.

Davis, P. M., & Fromerth, M. J. (2005). *Does the arXiv Lead to Higher Citations and Reduce Publisher Downloads for Mathematics Articles?*arXiv.org computer science. Retrieved 2/37/2006 from http://arxiv.org/abs/cs.DL/0603056.

Foster, A. (2006). Library leaders press colleges to archive online journals *Chronicle of Higher Education, 52*(26), A33.

Foster, N. F., & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine, 11*(1), [n].

Genoni, P. (2004). Content in institutional repositories: A collection management issue. *Library Management, 25*(6), 300-306.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage Publications.

Gibbons, S. (2005). *Institutional repositories listing (prepared for LITA regional institute, "establishing institutional repositories").* Retrieved 4/10, 2006 from http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-19825/IR%20Examples.pdf

Givler, P. (2002). University press publishing in the United States. In R. E. Abel, & L. W. Newlin (Eds.), *Scholarly publishing: Books, journals, publishers, and libraries in the twentieth century* (pp. 107-120). New York: Wiley.

Goodyear, M., & Fyffe, R. (2006). Institutional repositories: An opportunity for CIO campus impact. [Electronic version]. *Educause Review, 41*(2), 8-9. Retrieved 3/16/2006 from http://www.educause.edu/apps/er/erm06/erm0626.asp.

Harnad, S. (2003). Open access to peer-reviewed research through Author/Institution self-archiving: maximizing research impact by maximizing online access. In D. Law, & J. Andrews (Eds.), *Digital libraries: Policy, planning, and practice*. Burlington, VT: Ashgate Publishing.

Harnad, S. (1995). Overture: The subversive proposal. In A. S. Okerson, & J. J. O'Donnell (Eds.), *Scholarly journals at the crossroads: A subversive proposal for electronic publishing*. Washington D.C.: Office of Scientific & Academic Publishing, Association of Research Libraries.

Hawkins, B. L. (2006). Advancing scholarship and intellectual productivity: An interview with Clifford A. Lynch (part 1). [Electronic version]. *Educause Review, 41*(2) Retrieved 3/27/2006 from http://www.educause.edu/apps/er/erm06/erm0622.asp?bhcp=1

Hirtle, P. (2001). Editorial: OAI and OAS: What's in a name? *D-Lib Magazine, 7*(4), [n].

Houghton, J. W., with Steele, C., & Henty, M. (2003). *Changing research practices in the digital information and communication environment*. Canberra, Australia: Department of Education, Science, and Training. Retrieved February 19, 2006,

Jantz, R., & Giarlo, M. J. (2005). Digital preservation: Architecture and technology for trusted digital repositories. *D-Lib Magazine, 11*(6), [n].

Johnson, R. K. (2002). Institutional repositories: Partnering with faculty to enhance scholarly communication. *D-Lib Magazine, 8*(11), [n].

Joint Information Systems Committee (UK). (2005). *Digital repositories: Helping universities and colleges.* Retrieved 8/24/2005 from http://www.jisc.ac.uk/index.cfm?name=pub_repositories

Jones, P. (2006). Future library collections session. Conference of the Library Association of the University of North Carolina at Chapel Hill, March 13, 2006, Chapel Hill, NC.

Kotler, P., & Levy, S. J. (1969). Broadening the concept of marketing. [Electronic version]. *Journal of Marketing, 33*(1). Retrieved 3/17/2006, from Business Source Premier database.

Kranich, N. (2004). *The information commons: A public policy report*. New York University School of Law: Free Expression Policy Project. Retrieved 4/2/2006 from http://www.fepproject.org/policyreports/InformationCommons.pdf

Lavoie, B. F. (2003). *The incentives to preserve digital materials: Roles, scenarios, and economic decision-making.* Retrieved 6/20/2005 from http://www.oclc.org/research/projects/digipres/incentives-dp.pdf

Lawrence, S. (2001). Free online availability substantially increases a paper's impact *Nature, 411*, 521.

Lenhart, A., Horrigan, J., & Fallows, D. (2004). *Content creation online.* Washington, DC: Pew Internet and American Life Project. Retrieved 3/26/2006 from http://www.pewinternet.org/pdfs/PIP_Teens_Content_Creation.pdf

Lenhart, A., Simon, M., & Graziano, M. (2001). *The internet and education: Findings of the Pew Internet & American Life Project*. Washington, DC: Pew Internet and American Life Project. Retrieved 3/24/2006 from http://www.pewinternet.org/pdfs/PIP_Schools_Report.pdf.

Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy, 3*(2), 327-336.

Lynch, C. A., & Lippincott, J. K. (2005). Institutional repository deployment in the united states as of early 2005. *D-Lib Magazine, 11*(9), [n].

Mackie, M. (2004a). Filling institutional repositories: Practical strategies from the DAEDALUS project. [Electronic version]. *Ariadne,* (39)

Mackie, M. (2004b). The DAEDALUS project: Developing institutional repositories at Glasgow University. *Assignation, 21*(3), 16-18.

McClendon, W. (2005). *Institutional repositories: A white paper for the UNC-Chapel Hill Scholarly Communications Convocation*. Chapel Hill, NC: University of North Carolina. Retrieved 2/22/2006 from http://www.unc.edu/scholcomdig/whitepapers/mclendon.pdf

Miller, P. (2005). Web 2.0: Building the new library. [Electronic version]. *Ariadne,* (45)

Moeser, J. (2005). *State of the University Address*. University of North Carolina at Chapel Hill. Retrieved 3/27/2006 from http://www.unc.edu/chan/speech_archive/05stateofuniv.html

Montgomery, C. H. (2000). Measuring the impact of an electronic journal collection on library costs: A framework and preliminary observations [Electronic version]. *D-Lib Magazine, 6*(10), [n].

National Institutes of Health. (2006). *Report on the NIH public access policy.* Retrieved 4/1/2006 from http://publicaccess.nih.gov/Final_Report_20060201.pdf

National Institutes of Health. (2005). *Policy on enhancing public access to archived publications resulting from NIH-funded research.* Retrieved 4/1/2006 from http://publicaccess.nih.gov/

National Science Foundation. (2005). Dissemination and sharing of research results. *Grant policy manual.* Retrieved 4/2/2006 from http://www.nsf.gov/pubs/manuals/gpm05_131/gpm7.jsp#734

Nitterhouse, D. (2005). *Digital Production Strategies for Scholarly Publishers: A White Paper.* BiblioVault/University of Chicago. Retrieved 3/27/2006 from http://www.bibliovault.org/digital_prod_strategies.pdf

OCLC. (2003). *OCLC environmental scan 2003: Research and learning landscape.* Retrieved November 15, 2005 from www.oclc.org/membership/escan/research/repositories.htm

Panitch, J. M., & Michalak, S. (2005). *The Serials Crisis: A White Paper for the UNC-Chapel Hill Scholarly Communications Convocation.* Chapel Hill, NC: University of North Carolina. Retrieved 3/17/2006 from http://www.unc.edu/scholcomdig/whitepapers/panitch-michalak.html

Pinfield, S., & James, H. (2003). The digital preservation of e-prints. *D-Lib Magazine, 9*(9), [n].

Pomerantz, J., & Blouin, B. (2005). *Valuing Non-Traditional Vehicles of Scholarship: A White Paper for the UNC-Chapel Hill Scholarly Communications Convocation*. Chapel Hill, NC: University of North Carolina. Retrieved 3/14/2006 from http://www.unc.edu/scholcomdig/whitepapers/pomerantz-blouin.pdf

Prosser, D. C. (2003). Scholarly communication in the 21st century - the impact of new technologies and models. *Serials, 16*(2), 163-167.

Public Library of Science. *About PLoS: Mission and goals.* Retrieved 3/31/2006 from http://www.plos.org/about/index.html

Regier, W. G. (2003). Five problems and nine solutions for university presses. *Chronicle of Higher Education, 49*(40), B7.

Research Councils UK. (2006). *Summary of RCUK position statement on access to research outputs.* Retrieved 4/1/2006 from http://www.rcuk.ac.uk/access/summary.pdf

Research Libraries Group. (2002). *Trusted digital repositories: Attributes and responsibilities an RLG-OCLC report*. Mountain View, California: RLG, Inc. Retrieved 6/15/2005 from http://www.rlg.org.libproxy.lib.unc.edu/legacy/longterm/repositories.pdf

Richardson, M. (2005). Open access and institutional repositories: An evidence-based approach. *Serials, 18*(2), 98-103.

Stout, L. J. (1995). The role of university archives in the campus information environment. *The American Archivist, 58*, 124-140.

Stutzman, F. (2006). *Student life on the Facebook.* Retrieved 4/8/2006 from http://chimprawk.blogspot.com/2006/01/student-life-on-facebook.html

Suber, P. (2004). *Very brief introduction to open access.* Retrieved 4/1/2006 from http://www.earlham.edu/~peters/fos/brief.htm

Swan, A., & Brown, S. (2005). *Open access self-archiving: An author study*. Truro, Cornwall, UK: Key Perspectives Limited. Retrieved 2/20/2006 from http://www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

Tennant, R. (2002). Institutional repositories. *Library Journal, 127*(15), 28-30.

van Westrienen, G., & Lynch, C. A. (2005). Academic institutional repositories: Deployment status in 13 nations as of mid 2005. *D-Lib Magazine, 11*(9), [n].

Ware, M. (2004). *Publisher and Library/Learning solutions (PALS) pathfinder research on web-based repositories*. Bristol, UK: Publisher and Library/Learning Solutions. Retrieved 11/23/2005 from http://www.palsgroup.org.uk/palsweb/palsweb.nsf/0/8c43ce800a9c67cd80256e370051e88a/$FILE/PALS%20report%20on%20Institutional%20Repositories.pdf

Waters, Donald J., editor. (2005). *Urgent Action Needed to Preserve Scholarly Electronic Journals*, Digital Library Federation. Retrieved 3/14/2006 from http://www.diglib.org/pubs/waters051015.htm

Wilson, T. (2006). Institutional open archives: Where are we now? *Library and Information Update*, *5*(4), 31-33.